LWG – Chemical Fate and Transport Model
Initial Response to EPA Comments on 5/4/10 Presentation Materials

This document provides initial responses to EPA's comments received in a letter dated May 18, 2010 regarding the May 4, 2010 QEAFate model presentation. Per EPA's request, to support the agreed- upon upcoming small follow-up meeting to discuss the model, these initial responses focus on clarifying questions that the LWG has and an assessment of the work involved in replying to some of the additional information requests in these comments. As such, the LWG is not indicating at this time any agreement or disagreement with the comments. Each EPA comment is shown below in its entirety, followed by an initial response shown in blue.

EPA Comments and Direction on How to Proceed:
EPA is providing the following set of comments and questions that must be discussed before we can provide the approval to proceed with the QEAFate modeling approach as presented during our meeting on May 4, 2010. Due to the detailed nature of some of these comments and questions, EPA recommends a small follow-up meeting between the LWG modelers and the members of the EPA project team responsible for reviewing the Portland Harbor fate and transport model in order to get down to the fine details of their model and modeling procedures.

Although EPA is prepared to provide the go ahead for the modeling approach, EPA expects model results to have significant uncertainty no matter how good the calibration and validation is. As a result of this uncertainty, the model should be used in the FS in a comparative manner (i.e., evaluate the long term contaminant reductions for one alternative in comparison to another) and that the model results represent one line of evidence in the evaluation of monitored natural recovery at the Portland Harbor site and that other information such as grain size distributions, empirical measurements of bathymetric change, and any observed reductions in contaminant concentrations over the life of the project will also be considered.

Response: The LWG agrees that it would be good to have a follow-up meeting to discuss these comments as there are several that require some additional clarification. We see a key meeting objective being to determine the bare minimum list of additional information EPA needs now in order for EPA to provide approval to proceed with use of the model in the alternatives screening evaluation. This is as opposed to items that LWG could conduct as part of future modeling evaluations and present at the next check in.

With regard to model uncertainty, the LWG agrees that this model (not unlike most models) will have significant uncertainty no matter how good the calibration is. However, it is still a better tool than having no model at all. Without a model such as this, we are left to rely on the conceptual models, which are not always constrained by the principles of mass and energy balance. The key to dealing with model uncertainty will be to quantify this uncertainty, and factor this in to the model forecasting that will be performed during the FS. We also agree that the model only represents one line of evidence that will be included in the FS evaluation of monitored natural recovery at the Portland Harbor site.

Model Calibration:
The following information is needed to evaluate the adequacy of the model calibration. This information will assist EPA in deciding if the model is simulating the various physio-chemical processes correctly.

1. Plots of the average concentrations of the simulated contaminants over the top 15 cm of the sediment bed (in addition to those over the top 30 cm presented during the May 4, 2010 meeting) since that is most likely the concentration values that will be used in the food web modeling.

   Response: These plots are relatively easy to create.

2. Plots that show longitudinal and lateral (e.g., at select cross-sections) time series of concentrations over a multiple day period that includes a high flow event. This should be accompanied by a sufficiently detailed write-up that describes the simulated event and the model response over the course of the event (see attached figure).

   Response: These plots can be provided, however some additional clarification is needed. For example, is this request for all modeled contaminants? Also, is EPA requesting this analysis for the water column, sediment, or both? This analysis will take some time since analysis of a single high flow event will require re-running the model with outputs stored more frequently (current output interval is daily), particularly if EPA would like to see this for all modeled contaminants.

   LWG is also questioning the necessity of providing a detailed written explanation of the model response. This will take additional time to develop. It should be noted that there will be a detailed modeling report in the draft FS that will cover this and many of the other issues raised in these comments.

3. Explain how measured concentrations of contaminants in the sediment (particulate, DOC complexed, and freely dissolved phases) were used in calibrating the model.

   Response: We need clarification on specifically what is being requested.

4. Results from a process-based global mass balance analysis performed on a multi-year model run for each of the contaminants modeled. See the figure on the last page for an example of this type of analysis that was performed for a reach of the Housatonic River, MA.

   Response: The requested mass balance can be provided. This will take some time due to a few complexities associated with the time scales of analysis/averaging and the flow reversals that occur within the system. We intend to provide this in the detailed modeling report in the draft FS.

Additional Considerations:

1. Contaminant transport and fate model performance is best gauged by dynamic responses during high flow events and during base-flow as well rather than by long term temporal averages.

   Response: Although most of the material presented during the meeting focused on longer-term responses, we certainly looked at model response on shorter time scales (the animations provide one example). Because there really are not any high frequency data to assess model response, we would like to discuss with EPA how exactly this would be accomplished. Depending on their content, evaluations of smaller time scales could take considerable time.

2. Validation of the model was not discussed in detail during the May 4, 2010 meeting. The model will need to be validated before it is used to make multi-decade model runs to evaluate various proposed remedies. The validation strategy should be presented to EPA for review and comment.

   Response: Need to discuss. Our initial thought regarding model validation was that the data set was best utilized together to accomplish calibration, rather than parsing them into separate calibration and validation data sets. Further, there are not really any strong temporal patterns in the water column data to "validate" the model against if a portion of the data set had been held back for validation (i.e., we did not make any adjustments to the model during calibration to reproduce observed differences in the water column data collected near the beginning or end of the calibration period – there were no apparent temporal differences in the data over this 7-year period). If possible, a separate validation step would take considerable time.

3. It is unclear what time period the model output concentrations will be averaged over for the food web model. A range of averages (hourly to daily) should be provided to understand how much variation is occurring due to the influence of the tide in the lower reach of the Lower Willamette.

   Response: The current plan is to use daily averaged output from the fate model for the food web model. Fish uptake responds to longer-term average exposures. Short-term fluctuations in water column concentrations typically do not illicit a rapid response in biota levels (except possibly at the base of the food chain). Based on our experience in modeling bioaccumulation at several sites, we would expect the predicted fish levels to be the same whether exposures are provided on a daily or hourly timescale. Moreover, test simulations have been performed using the FWM (using daily- versus monthly-average exposure concentrations) that demonstrate little change in the FWM-predicted fish concentrations. Producing outputs on an hourly timescale would take considerable additional time.

4. EPA is concerned that the model calibration was based on the water data but the predicted reductions in sediment concentrations over the calibration period do not match up with what we observe in the sediment bed. This is clearly evidenced by looking at the scatter plots of sediment concentration vs. time and through consideration of key source areas at the site where contaminant concentrations remain high despite being released 30 – 50 years ago. The predicted water concentrations show a sharp, short-term increase in response to high flow events. This is not consistent with the surface water data collected during high flow events which actually see a decline in contaminant concentrations (although perhaps over a longer

time scale than is predicted by the model). Although EPA and the LWG agreed that we could not calibrate the model based on sediment because of spatial heterogeneity (in particular, we did not see any difference in paired sediment station concentrations whether the data were collected on the same day or 3000 days apart), further efforts to calibrate or validate the model using average sediment concentrations over time (e.g., 1 year) should be considered.

Response: We are currently in the process of reviewing the data to further explore sediment temporal trends. However there are a few specific aspects of this comment we need to discuss. We should discuss whether this evaluation can be conducted moving forward and not hold up the EPA approval to move to the next step.

5. There are a number of processes that need to be modeled. For example, the predicted water concentrations show a sharp spike in response to high flow events. What happens to these contaminants? Are they removed from the system? Do they settle right down? What happens to the more highly contaminated sediments below the eroded layer? Are we properly accounting for the mixing of these contaminants with the surface layer? Are sedimentation rates being properly estimated?

   Response: We should discuss whether these questions need to be answered now to obtain EPA approval for the next step or not. We should also discuss examples of the information that EPA would eventually accept as adequately answering each of these questions. This would affect the amount of time it would take to respond to EPA with answers. We also have some questions about the meaning of some of the questions.

6. Because the model is a two-dimensional model, the calibration assumes vertical mixing. However, based on surface water data collected at the site, near bottom surface water samples are expected to have higher contaminant concentrations. The model evaluation should consider the uncertainty associated with the use of a 2-dimensional model and whether there are ways to compensate for this assumption.

   Response: The significance of this is a matter of scale, since the time/distance over which vertical mixing in this system occurs is likely much smaller than the scales of importance for the fate and transport model. For example, the general equations that describe vertical mixing within a river (Fischer et al., 1979) indicate that a contaminant entering a stream (either near surface or near bottom) typically becomes completely mixed vertically at a distance downstream of approximately 10 to 20 times the channel depth.[1] Using an average water depth over the study area of approximately 12 meters, the water column would be completely mixed vertically at a distance of approximately 120 to 240 meters downstream. This distance is insignificant considering the length of a single model grid cell is approximately 180 meters.

   Given the above discussion, we feel the uncertainty associated with the use of a 2-D model is insignificant compared to other sources of uncertainty and does not need to be considered further. In any event, the decision to develop a 2-D versus 3-D model was agreed upon back when the HST was designed, and EPA subsequently agreed to consider the QEA Fate model which is supported by the 2-D HST model. We are unclear whether EPA is asking for additional uncertainty analyses or discussions before we can proceed to the next modeling step. We could

[1] *Fischer, H.B., List, J.E., Koh, R., Imberger, J., and Brooks, N.H. 1979. Mixing in Inland and Coastal Waters.*

add discussions of the uncertainty created by this assumption in the draft FS similar to the one provided above.

7. The model should be used to generate contaminant concentrations averaged over a ½ mile interval and divided laterally into two near shore areas and a channel area. Although smaller spatial scales may need to be evaluated, this should be performed as part of a recontamination analysis. For example, what is the recontamination potential associated with stormwater discharges to a small dredged area.

   Response: It is possible to generate model outputs on the smaller spatial scale requested. With regard to the evaluation of recontamination potential, we feel that this is something that would be done as part of the FS, and should therefore not hold up approval of the model.

8. Please clarify the relative importance of groundwater flux vs. contaminant concentrations in the plume. Slides 186-188 indicate that contaminant concentrations of naphthalene are essentially irrelevant and the difference in chemical concentrations in surface water is due to changes in the groundwater flow rate. Slide 209 states that groundwater flow is insignificant and that groundwater concentrations are insignificant with the exception of naphthalene. Additional information regarding the estimate of flux calculations should be provided.

   Response: We agree that Slide 209 is a bit confusing. The groundwater flow affects model results much more than concentration, although both did not strongly affect the calibration results except for naphthalene. We need clarification on the "additional information" that is requested in the last sentence of this comment and confirmation of whether this is needed to gain approval to move to the next modeling step. The time it would take to provide the information is dependent on its content.

9. Many of the model runs show concentration decreases near the beginning of the simulation and then appear to approach a new steady state in the last few years of the simulation. This has not been observed in the existing data (i.e. over the calibration time period 2002-2008) and needs to be explained further or substantiated with additional data. Depending on whether the new steady state concentration is less than or greater than the RAL, this could either indicate that natural recovery works very fast and would be relied on for a significant part of the remedy, or that the concentration flat-lines above the RAL due to ongoing inputs and thus remediation should not yet be attempted until further source control is achieved. Longer model runs would help clarify whether there is actually a new steady state or if this is just the result of flow conditions or other factors specific to the years 2006-2008.

   Response: See response to #4 above – need to discuss, however we are currently in the process of reviewing the data to further explore sediment temporal trends. We also agree that longer runs are needed to fully assess model-predicted sediment temporal trends. Such runs are planned as part of the FS simulations.

10. The observed contaminant concentration trajectory may represent overestimates in the model initial conditions, which gradually become less influential as other forcing functions have their effect. Or it could represent an over estimate of the biodegradation rate or some other contaminant removal parameter such that steady state conditions reflect only the ongoing lateral and upstream inputs. It may be possible to distinguish between these two possibilities by

comparing the predicted 2008 concentrations to the model inputs for upstream and lateral loads (lateral would include groundwater and stormwater and NPDES). If the "steady state" number is close to that, then it's probably case 2. Alternatively, we could look at the spatial coverage of the data set used to create the sediment bed initial conditions. Perhaps it's biased towards suspected/known contaminated areas and it's giving us an overly high initial value.

Response: See response to #4 above – need to discuss, however we are currently in the process of reviewing the data to further explore sediment temporal trends.

11. Data collection to support analysis of temporal trends was not done for the RI, but there may be sufficient spatial coverage at least for some of the chemicals (PAHs, for example). This would help clarify whether the "steady state" prediction is reasonable.

    Response: See response to #4 above – need to discuss, however we are currently in the process of reviewing the data to further explore sediment temporal trends.

12. It is unclear why PCB 126 and TBT were chosen for the sensitivity analysis; the data are relatively sparse for both these chemicals. It's not clear why PCB 126 was used and not a homolog group, since that is what is being modeled. The lack of data in some areas can drive the SWAC and thus the initial conditions.

    Response: Just to clarify, it appears that this comment is referring to the case where TBT and PCB 126 were examined for the sensitivity analysis on deep sediment (1-4 ft) initial conditions. The reason this sensitivity analysis was limited to these two chemicals is that the main objective of the analysis was to answer the question: how are the model predictions affected by uncertainty associated with a sparse subsurface data set. TBT and PCB congeners were the only contaminants which were judged to fall into this category

13. Water column animations were presented during the May 4, 2010 meeting. However, surface sediment animations should also be developed.

    Response: We have developed these for all runs but did not show them at the meeting. The files can be provided easily.

14. For metals, the model is under predicting water concentrations and over predicting sediment concentrations. This suggests a problem with the Kd used to partition the metals concentration.

    Response: Need to discuss. It is not clear what the basis is for saying the model over-predicts sediment concentrations for metals. If this comment is referring to the over-prediction of concentrations measured in particulate matter collected in sediment traps, this is correct. However, it should be noted that for copper for example, the model reproduces concentrations on particulates suspended in the water column quite well.

    As discussed during the meeting, we recognize this issue; we could not calibrate to concentrations in surface sediment, sediment traps, and water column particulates using a single Kd. This could be due to the simplistic representation of metals partitioning with a single Kd. We might consider using different Kd's in the sediments versus water column (to reflect differing redox conditions / pH) to improve this. If such differences are a factor, some of the discrepancy

may be related to how model results are compared to the sediment trap data because the data show significant differences between the sediment traps and the water column particulates (and we use the same model output to compare to both data sets [i.e., predicted concentrations of water column particulate matter]). However, this may end up being a limitation we need to live with for metals due to the simplified partitioning approach (i.e., the use of a simple Kd to represent a very complex process of partitioning in metals).

15. During the meeting, it was suggested that chemicals with "good" calibration be used to predict future contaminant concentrations and chemicals with "fair" calibration to evaluate comparatively. As we have stated previously, the goal of the model is to perform a comparative evaluation of remedial action alternatives in the FS. Further discussion is needed to understand how the predicted future sediment concentrations will be used in the remedy selection process.

    Response: We agree that comparative evaluations should be the focus for all contaminants. However, the model, despite its uncertainties, remains the best tool to develop absolute predictions of metrics such as when a certain PRG may (or may not) be met. Of course, we would need to recognize and quantify uncertainties associated with such use of the model; there are things we can do to quantify this uncertainty.